

Peierls-Nabarro Barrier and Protein Loop Propagation

Adam K. Sieradzan,^{1,2,*} Antti Niemi,^{1,3,4,†} and Xubiao Peng^{1,‡}

¹*Department of Physics and Astronomy, Uppsala University,
Ångströmlaboratoriet, Lägerhyddsvägen 1, 751 20 Uppsala, Sweden*

²*Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-952 Gdańsk, Poland*

³*Laboratoire de Mathématiques et Physique Théorique CNRS UMR 6083,
Fédération Denis Poisson, Université de Tours, Parc de Grandmont, F37200 Tours, France*

⁴*Department of Physics, Beijing Institute of Technology,
Haidian District, Beijing 100081, Peoples Republic of China*

When a self-localized quasiparticle excitation propagates along a discrete one dimensional lattice, it becomes subject to a dissipation that converts the kinetic energy into lattice vibrations. Eventually the kinetic energy does no longer enable the excitation to cross over the minimum energy barrier between neighboring sites, and the excitation becomes localized within a lattice cell. In the case of a protein, the lattice structure consists of the C α backbone. The self-localized quasiparticle excitation is the elemental building block of loops. It can be modeled by a kink which solves a variant of the discrete non-linear Schrödinger equation (DNLS). We study the propagation of such a kink in the case of protein G related albumin-binding domain, using the UNRES coarse-grained molecular dynamics force field. We estimate the height of the energy barriers the kink needs to cross over, in order to propagate along the backbone lattice. We analyse how these barriers gives rise to both stresses and reliefs which control the kink movement. For this, we deform a natively folded protein structure by parallel translating the kink along the backbone away from its native position. We release the transposed kink, and we follow how it propagates along the backbone towards the native location. We observe that the dissipative forces which are exerted on the kink by the various energy barriers, have a pivotal role in determining how a protein folds towards its native state.

PACS numbers: 36.20.-r, 87.15.B-, 36.20.Ey

I. INTRODUCTION

The protein folding problem has been around for a long time [1], with various incarnations [2–4]. Numerous proposals have been presented, to explain how the folding patterns emerge and advance. Examples include nucleation-condensation [5], hydrophobic collapse [6] and diffusion-collision controlled processes [7, 8]. More recently, a kink-based approach has been proposed [9, 10] and tested successfully using a coarse-grained force-field [11, 12]. At the same time, protein misfolding continues to be vigorously investigated [13]. A misfolded protein can lead to serious complications, a misfolding transition is now widely accepted as the cause for diseases like type-II diabetes, many forms of cancer, and dementia including Alzheimer’s [14].

In the present article we investigate the phenomenon of *barrier penetration* as an example of those physical mechanisms that contribute to protein folding and misfolding transitions. For this we inquire what happens when a protein loop fragment becomes displaced from its native position. How does it find its way back to the native position? The understanding of the dynamical processes involved should contribute to our comprehension of protein folding, including the reasons why and how a misfolding

transition takes place.

As a concrete example we consider an albumin-binding domain which is related to the protein G. This is a fairly generic α -helical protein, the native state involves three helices which are separated by two short loops; we use the Protein Data Bank (PDB) [15] entry code 1GAB. It consists of 53 amino acids, the first loop is located between sites 19-24 and the second loop is located between sites 33-37. The reason for choosing 1GAB as our concrete example is, that the UNRES force field, on which our investigations are based, has been optimised for this particular protein.

In our analysis of 1GAB we first follow [9, 10], to conclude that each of the two loops of 1GAB are modeled by a single kink that solves a generalized discrete non-linear Schrödinger equation [16–18]. We then parallel translate the second loop in the PDB structure by five residues towards the first loop, with no change in the remaining residues. We release the loop so that it is free to move along the backbone. This instigates a process where the kink attempts to propagate back to its original posture. Thus we have a prototype example of the Peierls-Nabarro scenario where a coherent and self-localized quasiparticle excitation propagates along a discrete one dimensional lattice [19–21]. Now realised in the case of the protein backbone.

*Electronic address: adasko@sun1.chem.univ.gda.pl

†Electronic address: Antti.Niemi@physics.uu.se

‡Electronic address: xubiaopeng@gmail.com

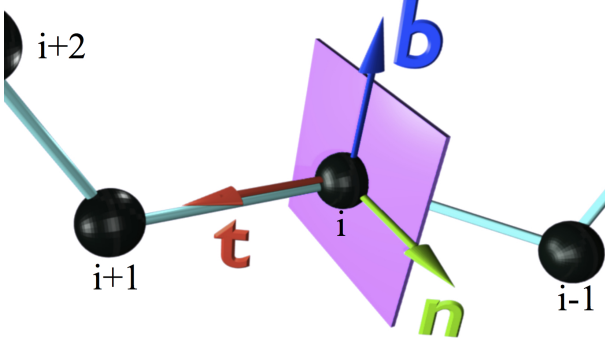


FIG. 1: (Color online) Frenet frames along the C^α backbone lattice.

II. METHODS

Our analysis is based on the United Residue (UNRES) force field [22–29]. It is freely available at

$$\text{http} : // \text{www.unres.pl} \quad (1)$$

UNRES is a carefully grafted coarse-grained force field that can be utilized to simulate the time evolution of a displaced kink along the protein backbone, in a highly realistic manner. UNRES has a very strong predictive power [30]. It has also been optimised to reproduce the correct thermodynamic properties, in a scenario like the one that will be studied here. Moreover, UNRES has a proven track record of already successfully modelling the kink formation in an α -helical protein [12], like the one studied here. With the presently available computer power, a coarse-grain force-field such as UNRES enables us to produce numerous full-length trajectories to probe the dynamical details in terms of a statistical analysis, which we needed in particular when determining the temperature dependence of the Peierls-Nabarro barrier.

A. Protein backbone geometry

We describe the skeletal C^α backbone lattice, using the geometrically determined Frenet frames [31]. These frames depends *only* on the position of the C^α carbons, with coordinates \mathbf{r}_i where $i = 1, \dots, n$ labels the residues. At a given residue, the orthonormal Frenet frame consists of the backbone tangent (\mathbf{t}), binormal (\mathbf{b}), and normal (\mathbf{n}) vectors. These are defined by equations (2), (3), and (4), respectively, and shown in figure 1.

$$\mathbf{t}_i = \frac{\mathbf{r}_{i+1} - \mathbf{r}_i}{|\mathbf{r}_{i+1} - \mathbf{r}_i|} \quad (2)$$

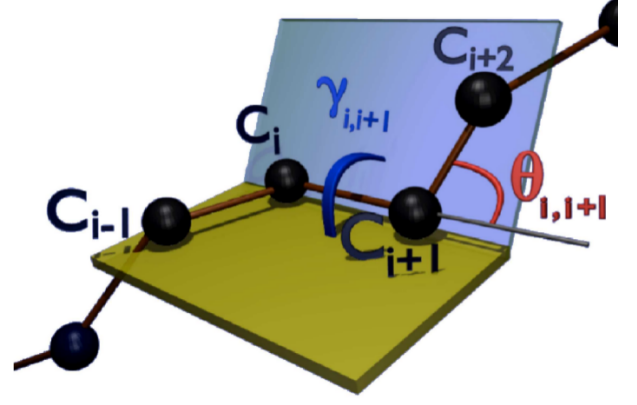


FIG. 2: (Color online) Definition of bond (θ_i) and torsion (γ_i) angles, along the discrete C^α string.

$$\mathbf{b}_i = \frac{\mathbf{t}_{i-1} \times \mathbf{t}_i}{|\mathbf{t}_{i-1} \times \mathbf{t}_i|} \quad (3)$$

$$\mathbf{n}_i = \mathbf{b}_i \times \mathbf{t}_i \quad (4)$$

For *trans*-peptide groups the distance between the consecutive C^α atoms is close to the constant value

$$|\mathbf{r}_{i+1} - \mathbf{r}_i| = \Delta \approx 3.8 \text{ \AA}$$

Thus the backbone lattice becomes entirely described by virtual-bond-valence angles θ and virtual-bond-dihedral (torsion) angles γ ; see figure 2.

$$\theta_{i+1,i} \equiv \theta_i = \arccos(\mathbf{t}_{i+1} \cdot \mathbf{t}_i) \quad (5)$$

$$\gamma_{i+1,i} \equiv \gamma_i = \text{sgn}[(\mathbf{b}_i \times \mathbf{b}_{i+1}) \cdot \mathbf{t}_i] \arccos(\mathbf{b}_{i+1} \cdot \mathbf{b}_i) \quad (6)$$

The frame vectors at site \mathbf{r}_{i+1} are expressed in terms of the frame vectors at site \mathbf{r}_i and the angles θ and γ using a transfer matrix,

$$\begin{pmatrix} \mathbf{n}_{i+1} \\ \mathbf{b}_{i+1} \\ \mathbf{t}_{i+1} \end{pmatrix} = \begin{pmatrix} \cos \theta \cos \gamma & \cos \theta \sin \gamma & -\sin \theta \\ -\sin \gamma & \cos \gamma & 0 \\ \sin \theta \cos \gamma & \sin \theta \sin \gamma & \cos \theta \end{pmatrix}_{i+1,i} \begin{pmatrix} \mathbf{n}_i \\ \mathbf{b}_i \\ \mathbf{t}_i \end{pmatrix} \quad (7)$$

The C^α lattice is constructed as follows,

$$\mathbf{r}_k = \sum_{i=0}^{k-1} |\mathbf{r}_{i+1} - \mathbf{r}_i| \cdot \mathbf{t}_i \equiv \Delta \sum_{i=0}^{k-1} \mathbf{t}_i \quad (8)$$

The normal and binormal vectors ($\mathbf{n}_i, \mathbf{b}_i$) are both absent in equation (8), only the tangent vector \mathbf{t}_i appears. Therefore, if these two vectors are simultaneously rotated around the vector \mathbf{t}_i , they continue to constitute an equally good framing. In particular a frame rotation by π determines the discrete \mathbb{Z}_2 transformation

$$\begin{aligned} \gamma_i &\rightarrow \gamma_i - \pi \\ \theta_k &\rightarrow -\theta_k \end{aligned} \quad \text{for all } k \geq i \quad (9)$$

that leaves (8) intact. This \mathbb{Z}_2 transformation has been used extensively to analyze proteins [11, 12, 31–35] and will also be used here.

B. Kink and protein geometry

The virtual bond angles (5) and virtual torsion angles (6) of the C^α backbone are mutually correlated, by the atomic level interactions along the protein chain. We may choose these angles as the order parameters, that characterise the thermodynamical state of the protein in terms of its geometric composition. In [9–11, 36, 37] it has been shown that in the limit of slowly varying variables, the full all-atom thermodynamical free energy F can be expanded in terms of these order parameters, in the universal sense of Kadanoff and Wilson [38, 39], as follows:

$$F = - \sum_{i=1}^{N-1} 2\theta_{i+1}\theta_i + \sum_{i=1}^N \left\{ 2\theta_i^2 + \lambda(\theta_i^2 - m^2)^2 + \frac{q}{2}\theta_i^2\gamma_i^2 - p\gamma_i + \frac{r}{2}\gamma_i^2 \right\} + \dots \quad (10)$$

In equation (10) λ , q , p , r , and m depend on the atomic level physical properties and the chemical microstructure of the protein and its environment. In principle these parameters can be computed from this knowledge.

The kink that constitutes the elemental building block of a protein loop arises from (10) as follows: The angles γ_i are first determined as functions of the angles θ_i ,

$$\gamma_i[\theta] = \frac{u}{1 + v\theta_i^2} \quad (11)$$

with $u = p/r$ and $v = q/r$. The equation (11) is inserted into equation (10), to eliminate the γ_i in favor of the θ_i . This gives a set of equations (12) for the θ_i

$$\theta_{i+1} = 2\theta_i - \theta_{i-1} + \frac{dV[\theta]}{d\theta_i^2}\theta_i \quad (i = 1, \dots, N) \quad (12)$$

where $\theta_0 = \theta_{N+1} = 0$ and

$$V[\theta] = - \frac{u^2 r}{2(1 + v\theta^2)} - 2\lambda m^2 \theta^2 + \lambda \theta^4 \quad (13)$$

Here we recognize the structure of a generalized discrete non-linear Schrödinger (DNLS) equation [9, 10, 16–18] In the case of proteins, u is small and $\lambda m^2 > 0$ so that there are two local minima in $V[\theta]$, with $\theta = \pm\theta_0 \approx \pm m$. The kink is a solution of (12), (11) that interpolates between the two minima. The explicit form of the kink is not known to us, in terms of elementary functions. But it can be constructed numerically by following the iterative procedure of ref. [10]. Note that a ground state configuration where θ_i is constant with

$$\begin{cases} \theta_i \approx \frac{\pi}{2} \\ \gamma_i \approx 1 \end{cases} \quad (\text{radians}) \quad (14)$$

describes a right-handed α -helix, while

$$\begin{cases} \theta_i \approx 1 \\ \gamma_i \approx \pm\pi \end{cases} \quad (\text{radians}) \quad (15)$$

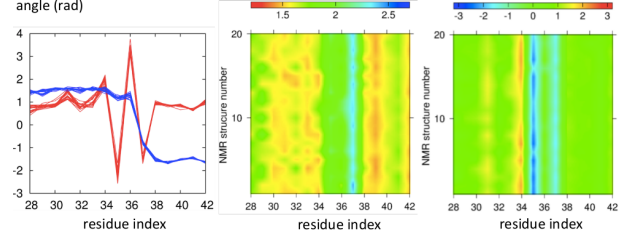


FIG. 3: (Color online) (Left) The θ (blue) and γ (red) profile of 1GAB, average over all NMR structures, after \mathbb{Z}_2 transformation. In (centre) and (right) we show the variations in θ and γ profiles, respectively, over the different NMR structures of 1GAB; colour intensity determines the value of the angle in radians, according to the scale on top.

describes a β -strand. For a kink, the (θ_i, γ_i) profile interpolates with these regular secondary structures.

For a given protein structure, we train the energy function (10) to describe the backbone geometry in terms of a multi-kink solution of (12), (11). For this we use the packages *Propro* and *GaugeIT*, described at

<http://www.folding-protein.org>

In the case of 1GAB, we use the averaged NMR structure. In figures 3 we describe the profile of the second loop along the 1GAB backbone, averaged over all NMR structures in PDB. In the figure 3 (left) we have implemented the \mathbb{Z}_2 transformation (9). It reveals that the loop is a single kink, that appears as a domain wall between two α -helical structures.

C. The Peierls-Nabarro barrier

In a protein, a loop propagation instigates a structural deformation. In crystallographic protein structures the loops are located at definite positions along the backbone chain. Thus, in the case of a static, crystallographic protein the translation invariance of the kink along the C^α lattice is clearly broken. The breakdown of translation invariance in the case of a lattice kink is in fact generic: Whenever a self-localized quasiparticle excitation such as a kink moves along the lattice, waves are emitted in its wake as vibrations in the lattice structure. These waves drain the kinetic energy of the quasiparticle excitation, causing it to decelerate. Eventually the kinetic energy becomes depleted to the extent that the excitation can no longer cross over the energy barriers between lattice cells, and instead it becomes localizes to a particular lattice cell.

Normally, a localized quasiparticle excitation has only two stationary positions relative to the lattice structure. Moreover, in a simple model these energy extrema are often positioned symmetrically, either at a lattice site or half-way between two neighboring lattice sites. One of

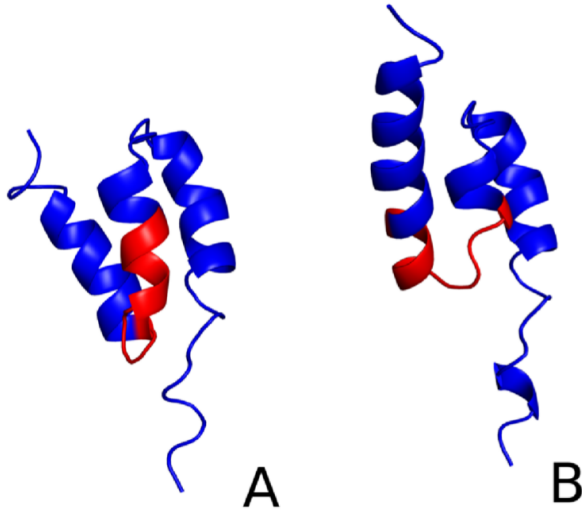


FIG. 4: (Color online) The protein G-related albumin-binding domain (PDB code: 1GAB) native structure (A), structure with shifted loop by five residues (B). The red fragment shows where the two structures differ.

the stationary positions is a local minimum of the potential energy, while the other is a saddle point. The energy difference between these two configurations determines the basal kinetic energy that the quasiparticle excitation needs, in order to translate itself from one lattice cell to another. This minimum energy difference is called the *Peierls-Nabarro barrier* (PNB) of the excitation [19–21].

D. Simulation details

We use the UNRES force field (1) in combination with the Molecular Dynamics (MD) protocol described in [40–42]. The initial condition for our simulations is the PDB structure 1GAB, where we have parallel transported the second loop fragment Ala³⁵-Val³⁸ to the position of Gln³⁰-Gln³³. Figure 4 compares the initial configuration with the crystallographic 1GAB structure. Note that the (θ_i, γ_i) spectrum remains intact during the parallel transport, except for the uniform loop translation. The ensuing C^α configuration is converted into an all-atom structure using a combination of Pulchra [43] and SCWRL4 [44]. Our choice to translate the loop fragment by five residues follows from an inspection of the all-atom steric hindrances that are incurred during the translation process.

We have simulated a total of 704 canonical molecular dynamics trajectories, at 11 different temperature values from 270K to 320K by steps of 5K. Thus there are 64 trajectories, at each temperature value. Each trajectory consists of 10 million iterations, with UNRES time step value 4.89 fs. Since the very fast atomic level motions are averaged by the UNRES force field, the corresponding *in vivo* time is around $\sim 50 \mu\text{s}$ [28, 41, 42].

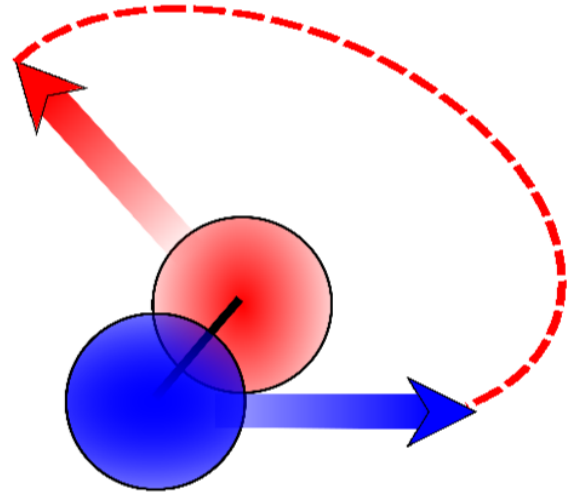


FIG. 5: (Color online) Schematic representation of the η_i angle. The sphere represents the C^α atom and the vector points toward the C^β position. The i^{th} residue site is marked with blue and the $(i + 1)^{\text{st}}$ is marked with red.

E. Analysis details

During the simulations we monitor the evolution of the (θ_i, γ_i) angles defined in (5) and (6), and the side-chain C^β-C^β torsional angles which we denote by η_i (schematically shown in the Figure 5). For this, the C^β positions are calculated by converting the coarse-grained UNRES trajectories into all-atom trajectories using an optimal dipole alignment algorithm developed in [45, 46].

We choose the 10th residue along the 1GAB backbone as a reference point: We have observed that it remain highly inert throughout our simulations. In particular, this residue does not seem to be affected by the kink propagation which we instigate. Starting with this reference site, we combine the side-chain C^β-C^β torsional angles η_i into the following accumulated *total* side-chain torsion angle $\hat{\eta}_i$,

$$\hat{\eta}_i = \sum_{k=10}^i \eta_k \quad (16)$$

During the simulations, we also monitor two quantitative criteria to assess whether the displaced kink has reached its native location. One of the criteria is the RMSD between the simulated structure and the PDB structure 1GAB. Following [47], we demand that for a native state, the RMSD should be less than 4Å. The second criteria is based on the (θ_i, γ_i) profile of the kink, as shown in figures 3. We demand that at the end, the kink is located within one lattice site from the native location between Ala³⁵-Val³⁸.

Finally, to estimate the height of the Peierls-Nabarro barrier, we evaluate the energy of the fragment Gln³⁰-Val³⁸ using the UNRES energy function.

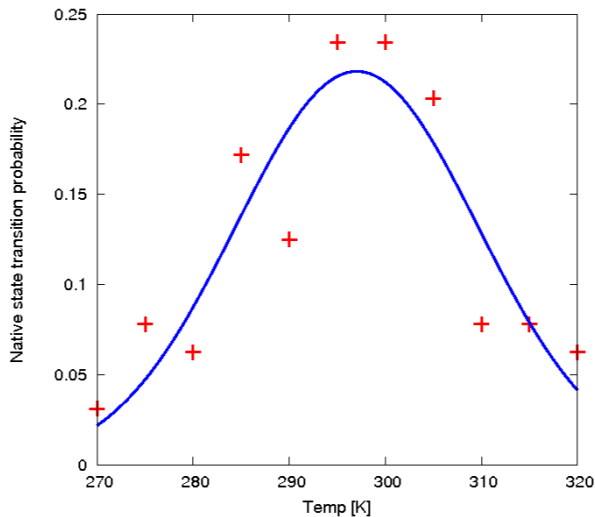


FIG. 6: (Color online) Fraction of simulations in which kink propagates to the native position, as a function of temperature. The red dots are the measured fraction, the blue line is a gaussian fit.

III. RESULTS AND DISCUSSION

A. Multiple pathways

According to [47], when the 1GAB folding is simulated at around 295K using UNRES force field with replica-exchange molecular dynamics, over 90% of structures reach the native state. However, in our UNRES simulations, using canonical molecular dynamics and *with our initial condition* where the second loop fragment has been parallel translated by five residues towards the first and then released at rest, we find that at 295K temperature only around 25 per cent of structures reach the native state. In figure 6 we show the probability for our initial configuration to reach the native state, as a function of temperature. The probability has a maximum at around 295K. We find that at low temperatures, well below 295 K, the Peierls-Nabarro energy barrier becomes too high for the kink to leap over it, only by the thermally induced kinetic energy. When the temperature increases and the amplitude of thermal fluctuations in the kink is larger, the barrier crossing becomes more and more frequent. But when temperature reaches values which are well above 295 K, the unfolding processes which are driven by increased thermal fluctuations in the protein lattice structure start becoming increasingly prominent and eventually the protein unfolds.

We also find that at temperatures well below 295 K, when the Peierls-Nabarro energy barrier becomes too high for the kink to cross over, the most frequently occurring pathway entails a disintegration of the kink; see Figure 7. Along this pathway, the parallel transported loop misfolds into a helical shape, and at the end we

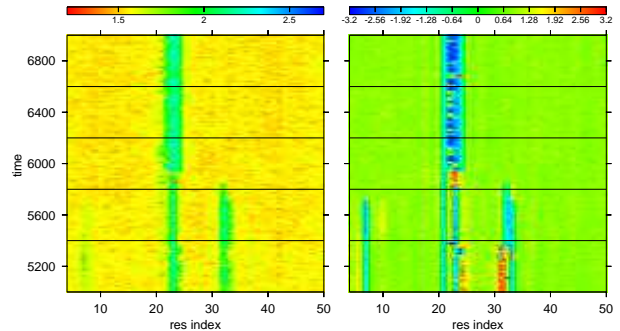


FIG. 7: (Color online) The changes of virtual valence angle θ (left) and torsional angle γ (right), as a function of simulation time and residue index during a process of kink disintegration.

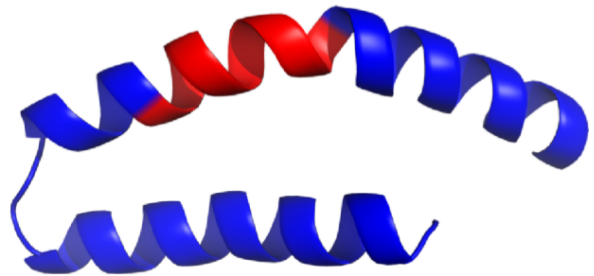


FIG. 8: (Color online) The structure of missfolded 1GAB protein after kink dissolution. The red color denotes the initial position of the parallel translated kink.

have a helix-bend-helix configuration akin the one shown in figure 8. We find that the energy barrier for this transition is quite low, as shown in figure 9.

B. Kink propagation

It should be noted that in our simulations each trajectory describes kink propagation in a somewhat different manner, no two trajectories are fully identical. There are also trajectories with kink-kink interactions. Only the general aspects of the most commonly occurring pathways are detailed in the sequel.

The kink propagation commences with a process of tension reduction: In the initial stage, the first and third helix are aligned in parallel, next to each other. Interactions between the ensuing side chains are present, causing a tension which hinders the propagation of the parallel transported kink. This tension becomes relaxed by a deformation that orients the first and third helix perpendicular to each other (see figure 10). This lowers the Peierls-Nabarro barrier, enabling the kink to start propagating.

Subsequently, we observe that the kink propagates along the chain by *reptation*: There is an initial pro-

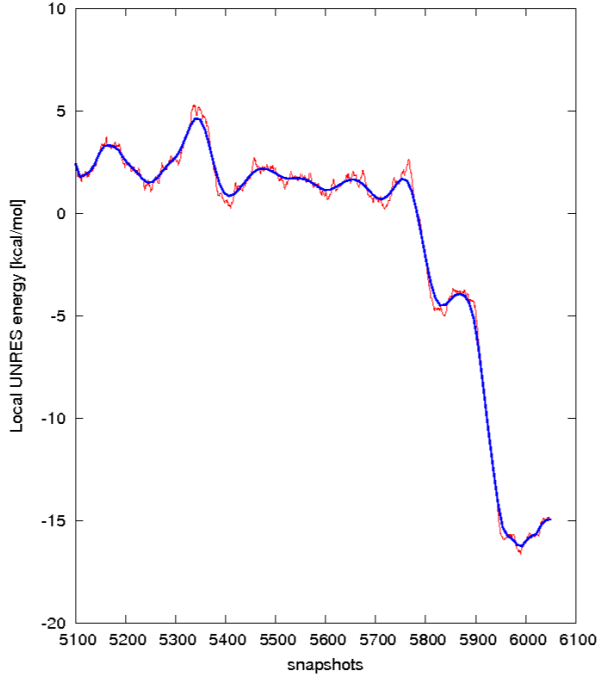


FIG. 9: (Color online) The local energy change during kink dissolution, evaluated over residues 29-40. The red line is the moving average over 50 snapshots and the blue line is a smoothed moving average.

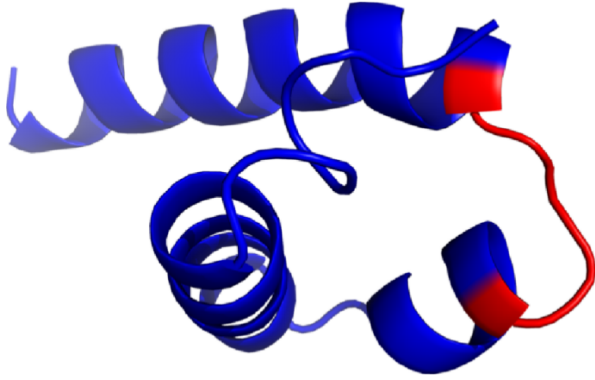


FIG. 10: (Color online) The structure of 1GAB protein after the initial transition with tension reduction. Note that the third helix becomes perpendicular to second helix which reduces the height of the Peierls-Nabarro energy barrier.

longation of the kink structure, an example is shown in figure 11. In this particular example the kink extends from residue 30 to residues 36-38, during the UNRES time period between ≈ 1000 and ≈ 3200 . The prolonged kink corresponds to an extended loop region, which is shown in figure 12; we also note that the third helix is perpendicular to both the first and the second helix. This enhances the conformational flexibility of the loop fragment which enables an increase in the accumulation

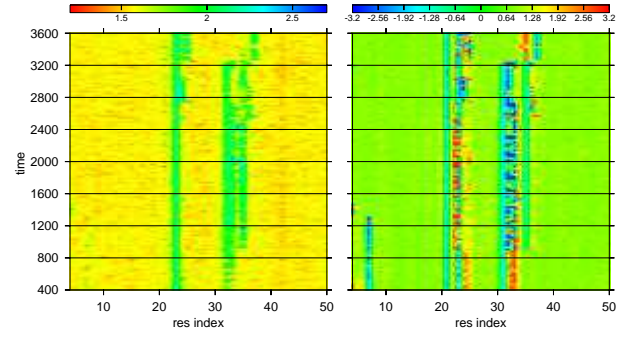


FIG. 11: (Color online) The variations of θ (left) and γ (right), as a function of time and residue index, during the kink propagation towards the native position; the color-coding is defined on top.

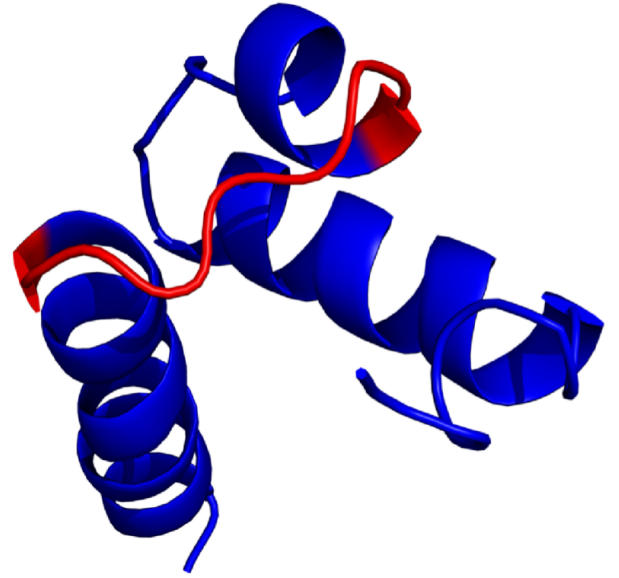


FIG. 12: (Color online) The structure of 1GAB protein when the kink is prolonged.

of thermally induced kinetic energy, boosting the kink to more easily cross over the energy barriers towards the final position. In the final stage the third helix then returns to a posture where it is (anti)parallel to the other two helices, and the kink assumes its native location.

C. Side-chain motion

The kink propagation is strongly correlated with the side-chain motions, as seen in the Frenet frames. For this, we first analyze the $C^\beta-C^\beta$ torsional angles η_i . The initial values of these angles, for the parallel transported kink, are shown in column A of figure 13. This figure shows that the value of η_i along a helical structure is $\eta_i \approx 0.8$ radians whereas along the kink the η_i deviate from this

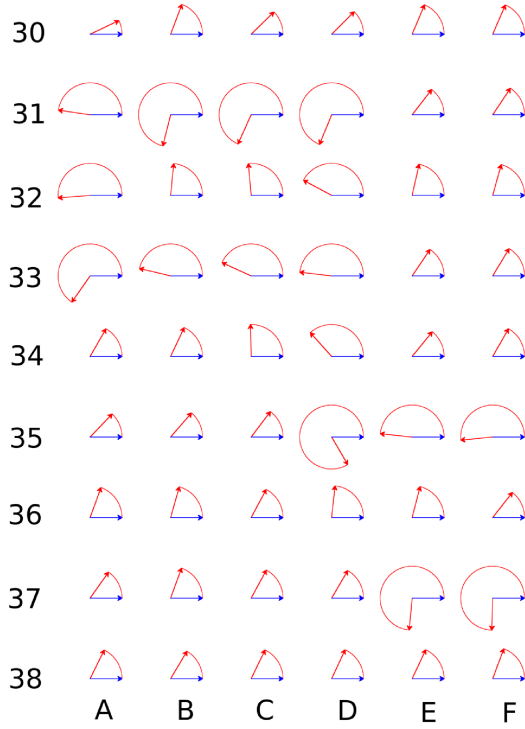


FIG. 13: (Color online) The times series (columns A-F) of η angles (residues 30-38).

helical value. In figure 13 column B we observe how the values of η_i remain largely intact during the initial phase, that of tension reduction, when the relative positions of the helices along the chain do not change.

During the second stage, that of reptation when the kink becomes prolonged, corresponding to the columns C and D in figure 13, the values of the η_i angle deviate from the helical value $\eta \approx 0.8$ radians, over an expanded range. There is a clear counter-clockwise relative rotation of the side-chains along the chain, correlating with the prolongation of the reptating kink along the chain.

In the final stage, as the third helix assumes its (anti)parallel orientation with the first and second helices, the length of the reptating kink becomes contracted into a region that extends over the residues 35-37 as the kink moves to its native position. This can be seen in column E and F of Figure 13.

Additional structure can be identified in terms of the accumulated total torsion angles $\hat{\eta}$, that we introduced in equation (16): Unlike the $C^\beta-C^\beta$ torsional angles η_i which are defined locally, the total angles $\hat{\eta}_i$ contain both local and global information. In figure 14 we show the values of $\hat{\eta}_i$ that correspond to the values of η_i in figure 13.

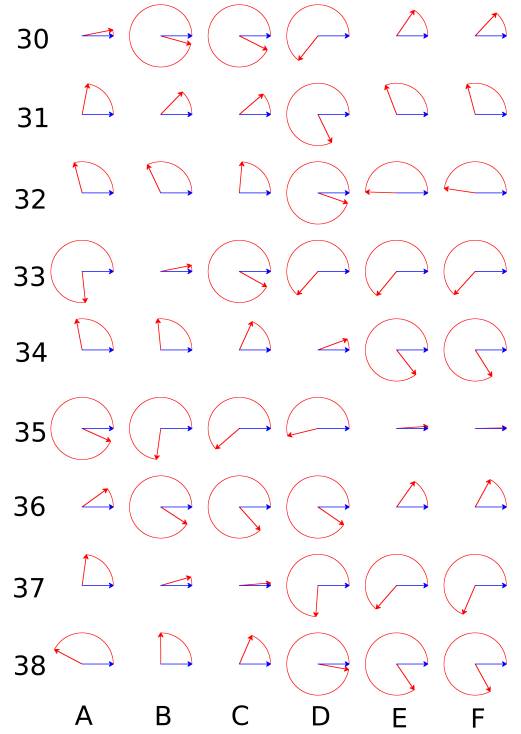


FIG. 14: (Color online) The times series (A-F) of $\hat{\eta}$ angles (residues 30-38).

For example, compare the first row (residue 30) in figures 13 and 14: In figure 13 there is no change, but in columns A-C of row 1 in figure 14 we see a substantial counter-clockwise rotation. Thus, we conclude that the kink propagation correlates with a global helical counter-clockwise rotation of the side-chains, that occurs prior to the position of the propagating kink.

Note that in both figures 13 and 14 there is no relative side-chain rotation visible between the two last columns. The value of η_{36} (angle between residue 36 and 37) remains essentially constant throughout our simulations. This orientational locking of the two side-chains with each other reflects a strong repelling interaction between the two lysines which are located at sites 35 and 37. This makes the two residues 36 and 37 to rotate in tandem.

D. Height of the Peierls-Nabarro barrier

We have monitored the change of local energy associated with the kink propagation; see figure 15. We observe that the difference in the energy which is localized at the initial parallel transported kink, and the energy which is localized at the final natively placed kink, is very

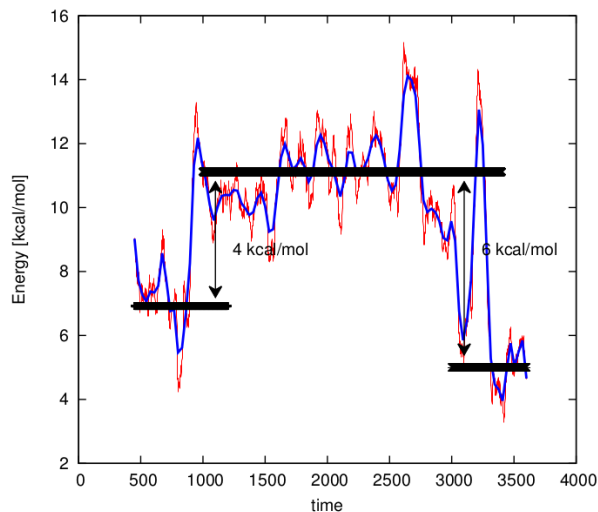


FIG. 15: (Color online) The local energy changes during the kink propagation, evaluated over the residues 29-40. The red line represents the moving average over the 50 snapshots and the blue line is a smoothed moving average.

small. Our estimate for the energy difference between the two kinks is around 1-2 kcal/mol. However, the energy barrier that the parallel transported kink needs to cross over in order to reach its native location, is much higher. From figure 15 we estimate that the height of the Peierls-Nabarro barrier is around 4-6 kcal/mol. This estimate is comparable, albeit slightly lower than the result obtained in [12]. It appears that the major contribution towards the height of the barrier correlates with the elongation of the kink, during reptation.

IV. SUMMARY

In this article we have analysed the propagation of a kink along protein backbone. The process is much more complex than a simple translation, it involves both local and global deformations. In particular, the Peierls-Nabarro barrier crossing involves a reptation-like movement, with the kink propagating by a combination of prolongations and contractions, and with a very sudden character. We note that this is in line with the analysis

of soliton propagation along a cyclic molecular chain [48].

We have estimated the height of the ensuing Peierls-Nabarro barrier to be around 4-6 kcal/mol, in the case of the second loop of 1GAB. We observe that this is remarkably close to the energy gap in ATP to ADP hydrolysis process.

It appears that the present study is the first to address the relevance of Peierls-Nabarro barrier to protein folding. Moreover, we have only considered a particular example. However, most protein loops can be described in terms of different parametrizations of the kink solution which is supported by (12), (11). Thus we trust that the phenomenon we have revealed is not limited to the specific example, but relates to a general mechanism how proteins fold and unfold.

Finally, we note that for a detailed investigation of loop propagation and the Peierls-Nabarro barrier, an all-atom molecular dynamics simulation should be performed for example along the lines of [49]. Such simulations, in particular over extended temperature range as described here remain highly time consuming, thus we postpone a detailed all-atom investigation to a future article.

V. ACKNOWLEDGMENTS

AKS research at Uppsala University was supported by Swedish Institute scholarship, AJN acknowledges support from Region Centre Recherche d'Initiative Academique grant, Sino-French Cai Yuanpei Exchange Program (Partenariat Hubert Curien), Vetenskapsrådet, Carl Trygger's Stiftelse för vetenskaplig forskning, and Qian Ren Grant at BIT. This research has been supported by an allocation of advanced computing resources provided by the National Science Foundation (<http://www.nics.tennessee.edu/>), and by the National Science Foundation through TeraGrid resources provided by the Pittsburgh Supercomputing Center. Computational resources were also provided by (a) the supercomputer resources at the Informatics Center of the Metropolitan Academic Network (IC MAN) in Gdańsk, (b) the 624-processor Beowulf cluster at the Baker Laboratory of Chemistry, Cornell University, and (c) our 184-processor Beowulf cluster at the Faculty of Chemistry, University of Gdańsk.

-
- [1] C. Anfinsen, H. Scheraga, *Adv. Protein Chem.* **29** 205 (1975)
 - [2] K.A. Dill, S.B. Ozkan, M.S. Shell, T.R. Weikl, *Ann. Rev. Biophys.* **37** 289 (2008)
 - [3] K.A. Dill, J.L. MacCallum, *Science* **338** 1042 (2012)
 - [4] B. Montgomery Pettitt, *Journ. Biomol. Struct. Dyn.* **31** 1024 (2013)
 - [5] E. López-Hernández, L. Serrano, *Folding and Design* **1** 43 (1996)
 - [6] R. Zhou, X. Huang, C.J. Margulis, B.J. Berne, *Science* **305** 1605 (2004)
 - [7] D. Bashford, F. Cohen, M. Karplus, I. Kuntz, D. Weaver, *Proteins* **4** 211 (1988)
 - [8] M. Karplus, D.L. Weaver, *Prot. Sci.* **3**, 650 (1994)
 - [9] M. Chernodub, S. Hu, A.J. Niemi, *Phys. Rev.* **E82** 011916 (2010)

- [10] N. Molkenhuth, S. Hu, A.J. Niemi, Phys. Rev. Lett. **106** 078102 (2011)
- [11] A. Krokhotin, A. Liwo, A.J. Niemi, H.A. Scheraga, J. Chem. Phys. **137** 035101 (2012)
- [12] A. Krokhotin, A. Liwo, G.G. Maisuradze, A.J. Niemi, H.A. Scheraga, J. Chem. Phys. **140**, 025101 (2014)
- [13] D. A. Butterfield, Free radical research **36** 1307 (2002)
- [14] C. Soto, E.M. Sigurdsson, L. Morelli, R.A. Kumar, E.M. Castaño, B. Frangione, Nature medicine **4** 822 (1998)
- [15] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, Nucl. Acid Res. **28** 235 (2000)
- [16] L.D. Faddeev, L.A. Takhtajan, *Hamiltonian Methods in the Theory of Solitons* (Springer-Verlag, Berlin, 2007)
- [17] M.J. Ablowitz, B. Prinari, A.D. Trubatch, *Discrete and continuous nonlinear Schrödinger systems* (Cambridge University Press, Cambridge, 2004)
- [18] P.G. Kevrekidis, *The Discrete Nonlinear Schrödinger Equation: Mathematical Analysis, Numerical Computations and Physical Perspectives* (Springer-Verlag, Berlin, 2009)
- [19] R. Peierls, Proc. Phys. Soc. **52** 34 (1940)
- [20] F. Nabarro, Proc. Phys. Soc. **59** 256 (1947)
- [21] F. Nabarro, Mat. Sci. Eng. **A234** 67 (1997)
- [22] A. Liwo, S. Oldziej, M.R. Pincus, R.J. Wawak, S. Rackovsky, H.A. Scheraga, J. Comput. Chem. **18** 849 (1997)
- [23] A. Liwo, M.R. Pincus, R.J. Wawak, S. Rackovsky, S. Oldziej, H.A. Scheraga, J. Comput. Chem. **18** 874 (1997)
- [24] A. Liwo, C. Czaplewski, J. Pillardy, H.A. Scheraga, J. Chem. Phys. **115** 2323 (2001)
- [25] A. Liwo, M. Khalili, C. Czaplewski, S. Kalinowski, S. Oldziej, K. Wachucik, H. Scheraga, J. Phys. Chem. **B111** 260 (2007)
- [26] A. Liwo, C. Czaplewski, S. Oldziej, A.V. Rojas, R. Kaźmierkiewicz, M. Makowski, R.K. Murarka, H.A. Scheraga, In *Coarse-Graining of Condensed Phase and Biomolecular Systems*; G. Voth, Ed.; (Taylor & Francis, 2008) Chapter 8, pp. 107-122.
- [27] U. Kozłowska, G.G. Maisuradze, A. Liwo, H.A. Scheraga, J. Comput. Chem. **31** 1154 (2010)
- [28] A.K. Sieradzan, H.A. Scheraga, A. Liwo, J. Chem. Theor. Comput. **8** 1334 (2012)
- [29] A.K. Sieradzan, U.H.E. Hansmann, H.A. Scheraga, A. Liwo, J. Chem. Theory Comput. **8** 4746 (2012)
- [30] Y. He, M.A. Mozolewska, P. Krupa, A.K. Sieradzan, T.K. Wirecki, A. Liwo, K. Kachlishvili, S. Rackovsky, D. Jagie, R. Iusarz, C.R. Czaplewski, S. Oldziej, H.A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **110** 14936 (2013)
- [31] S. Hu, M. Lundgren, A.J. Niemi, Phys. Rev. **E83** 061908 (2011)
- [32] M. Lundgren, A. Krokhotin, A.J. Niemi, Phys. Rev. **E88** 042709 (2013)
- [33] S. Hu, Y. Jiang, A.J. Niemi, Phys. Rev. **D87** 105011 (2013)
- [34] A. Krokhotin, M. Lundgren, A.J. Niemi, X. Peng Journ. Phys. Cond. Mat. **25** 325103 (2013)
- [35] A. Krokhotin, A.J. Niemi, X. Peng, Journ. Chem. Phys. **138** 175101 (2013)
- [36] A.J. Niemi, Phys. Rev. **D67** 106004 (2003)
- [37] U.H. Danielsson M. Lundgren, A.J. Niemi, Phys. Rev. **E82** 021910 (2010)
- [38] L.P. Kadanoff, Physic **2** 263 (1966)
- [39] K.G. Wilson, Phys. Rev. **B4** 3174 (1971)
- [40] M. Khalili, A. Liwo, F. Rakowski, P. Grochowski, H. Scheraga, J. Phys. Chem. **B109** 13785 (2005)
- [41] M. Khalili, A. Liwo, A. Jagielska, H. Scheraga, J. Phys. Chem. **B109** 13798 (2005)
- [42] A. Liwo, M. Khalili, H.A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **102** 2362 (2005)
- [43] P. Rotkiewicz, J. Skolnick, J. Comput. Chem. **29** 1460 (2008)
- [44] G.G. Krivov, M.V. Shapovalov, R.L. Dunbrack, Proteins **77** 778 (2009)
- [45] R. Kaźmierkiewicz, A. Liwo, H.A. Scheraga, J. Comput. Chem. **23** 715 (2002)
- [46] R. Kaźmierkiewicz, A. Liwo, H.A. Scheraga, Biophys. Chem. **100** 261 (2003) *Erratum:* Biophys. Chem. **106** 91 (2003).
- [47] P. Krupa, A.K. Sieradzan, S. Rackovsky, M. Baranowski, S. Oldziej, H.A. Scheraga, A. Liwo, C. Czaplewski, J. Chem. Theory Comput. **9** 4620 (2013)
- [48] L. Brizhik, A. Eremko, L. Cruzeiro-Hansson, Y. Olkhovska, Phys. Rev. **B61** 1129 (2000)
- [49] P. Cossio, A. Trovato, F. Pietrucci, F. Seno, A. Maritan, A. Laio, PLoS Comput. Biol. **6** e1000957 (2010)